DOCUMENT RESUME

ED 299 842 HE 021 220

AUTHOR	Fincher, Cameron
TITLE	Assessing Educational Outcomes: Are We Doing Good, Can We Do Better? IHE Newsletter.
INSTITUTION	Georgia Univ., Athens. Inst. of Higher Education.
PUB DATE	Feb 88
NOTE	8p.; Paper originally presented at the Annual Conference of the Mid-American Association for Institutional Researchers (Kansas City, MO, 1986).
AVAILABLE FROM	IHE Newsletter, Institute of Higher Education, University of Georgia, Athens, GA 30602.
PUB TYPE	Reports - General (140) -- Viewpoints (120) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Ability; *Academic Achievement; Basic Skills; *Educational Assessment; Educational Objectives; Higher Education; Measurement Techniques; Minimum Competency Testing; National Competency Tests; Needs Assessment; *Outcomes of Education; Public Policy; Student Characteristics
IDENTIFIERS	*College Outcomes Assessment; College Sophomores; *Competency Based Assessment

ABSTRACT

The United States is not doing as well as it could in assessing educational outcomes in the 1980s. The assessment movement of the 1980s indicates that its institutional memory is poor, and when coping with outside pressures, it is slow to recall how it coped the last time it was under pressure. It is important to remember such events and pressures as the minimal competency testing movement of the 1970s; the debates over the relative merits of criterion-referenced testing and norms-referenced testing; and the recurring love/hate affair that public education has with measurement, assessment, and evaluation. One reason the U.S. is not doing well in the assessment of educational outcomes is the confusion about educational purposes since the objectives and expected outcomes of higher education have not been defined.. Also, there is not an adequate theory of educational achievement in the 1980s. There is a need for a common language for education from kindergarten through "grade 16." The U.S. can learn from the experiences of the University System of Georgia (USGA) in creating systemwide entrance requirements, developmental studies programs, reading and writing tests for sophomores, and varying forms of senior exit exams. This system suggests that in any efforts to assess educational outcomes, there should be at least three stages of assessment. USGA has entry, rising junior, and senior exit assessment. Systematic, objective, valid, reliable, and fair measures of educational outcomes are rare. For the time being, college administrators should measure what they can, assess what they must, and evaluate with great care. Contains 15 references. (SM)

ED299842

ASSESSING EDUCATIONAL OUTCOMES: ARE WE DOING GOOD, CAN WE DO BETTER

by

Cameron Fincher

Institute of Higher Education
University of Georgia

HE021220

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Cameron Fincher

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ASSESSING EDUCATIONAL OUTCOMES: ARE WE DOING GOOD, CAN WE DO BETTER?

BY
CAMERON FINCHER

Institute of Higher Education
University of Georgia

The title to this commentary on assessment has been graciously supplied by Joe Saupe, University Director of Institutional Research at the University of Missouri.[1] We can answer his two questions in the following manner: Are we doing good? No, not really! Can we do better? We had better do better—or someone else will do the job for us!

The reason for both answers is simple. The assessment movement of the 1980s is but yet another indication that our institutional memories are poor and whenever we must cope with outside pressures, we are slow to recall how we coped the last time we were under pressure.

Events, movements, and pressures that should be remembered are: (1) the minimal competency testing movement of the 1970s, (2) the dreary debates over the relative merits of criterion-referenced testing and norms-referenced testing, and (3) the recurring love-hate affair that public education has with measurement, assessment, and evaluation. Because of our poor memories, we often forget that testing is one of the few genuine technological innovations we have had in education until the advent of xerography, personal computers, and video-cassette recorders. The development of multiple-choice tests, IBM answer sheets, and norms-referenced scoring is among the most underappreciated events in the history of education. Most of us remember only the periodic diatribes that excoriate standardized tests and appear briefly on best seller lists.

One of many reasons why we are not yet doing good in the assessment of educational outcomes is that we are much confused about educational purposes. We have not taken the time to define the objectives and expected outcomes of higher education, and much of what we do is lacking in adequate conceptual underpinnings. Our timidity should remind us of Bertrand Russell's statement that, "Many people would rather die than think; many do!"

But more importantly, in the 1980s we do not have an adequate or acceptable theory of educational achievement. We would like to measure academic competencies, but we do not want to use a test like the SAT because of group, age, and sex differences in verbal and mathematical abilities. And we are not certain that the SAT measures the kind of verbal and mathematical abilities that we ought to measure. Ignoring sixty years of research and development with the SAT, we play Hamlet and soliloquize about uses of the SAT in predicting freshman grades or in the decisions of admission officers. And yet, if intellectual development; the developed ability of students to reason with verbal and numerical concepts and symbols; or such matters as vocabulary, reading comprehension, and abstract reasoning are expected educational outcomes, the SAT is quite simply the best measure we have of such competencies in young adults.

In secondary education we seek tests of basic skills that all high school seniors can pass; in the first year of college we need tests of basic skills that will permit placement in developmental studies; and upon completion of lower division coursework we again measure basic skills in reading and writing to assure that students are capable of continuing their education. And in more recent years we have sought measures of similar basic skills for graduating seniors. Cynics may suspect that having established three check-points at the borders of literacy, we must establish a fourth to catch those with forged passports.

Those-in-charge assure those-of-us who are simple-minded that the basic skills measured in high school are different from the basic skills measured the following year in

[1] *. . . and Program Chair for the 1986 MIDAIR Conference in Kansas City where these remarks served as a keynote address.*

college—and completely different from the basic academic competencies identified in The College Board's Project EQuality (1983). And college officials patiently explain the need to measure basic skills of literacy at the "rising junior" level; faculty members who teach freshman and sophomore courses are apparently innocent of any reliable methods of assessing whether students can read and write, and academic standards simply must be imposed somewhere!

The challenge of "doing good" is nothing less than the challenge to define, assess, and teach academic competencies (and other educational outcomes) in ways that make educational sense. Our greatest need may be a common language for education from kindergarten through "Grade 16." As long as public school teachers talk about communication arts and college teachers lecture on literary nuances, students will continue to read and write only what they must—and as little as they can!

Granting the low probability of educational utopia, it is not too much to ask for a common language to discuss education in the last three years of high school and the first two years of college. These are the years which now appear most crucial to the development of reading, writing, and reasoning competencies. And despite the hegemony of state boards of education in Grades 10-12 and the supremacy of governing boards in Grades 13-14, many colleges and schools are now finding that they can work together on common educational problems. Optimism should be restrained, however, until high school teachers and college instructors have more experience in working together and until there are better methods of assessing reading, writing, and reasoning skills that are well defined and carefully taught.

Some of us would like to think that an acceptable taxonomy of educational objectives is emerging from current and past concerns with assessment. Howard Bowen (1977) and Mortimer Adler (1985) write in terms that are sufficiently similar for us to pay close attention. Whatever educational outcomes might be, some of us believe that they can be packed, without undue breakage, in the three boxes of knowledge, competence, and understanding. The resurgence of cognitive psychology in the past twenty-five years gives better significance to information-processing models of the human mind and underscores the importance of knowledge as an educational outcome. There is indeed a time and place in school to learn facts and figures, basic concepts, and general principles—and it is distinctly possible that we should seek to assess the learner's comprehension of such matters.

Professional education has long placed a premium on competence as expected outcomes, and much of education continues to be concerned with the development of skills and abilities that will be assessed in due time with carefully developed standards of performance. Competency-based education is often the focal point of educational attention, and in turn, our notions of competence, ability, and/or accomplishment have been quite influential in our efforts to assess educational outcomes. It is the apparent inability of high school and college graduates to demonstrate satisfactorily skills of literacy that has spurred much of our concern with assessment in the 1980s.

As most of us suspect, it is easier to assess the acquisition of knowledge than the development of competence—and both knowledge and competence would appear to be more amenable to measurement and assessment than the attitudes, beliefs, and values (our ABV's) so often expected as an outcome of education. Understanding is the most useful term to denote the benefits of personal experience (and is, little doubt, the expected outcome of experiential learning). The key to assessment here is evidently in the choices and decisions that educated individuals make. Personal interests, preferences, opinions, and/or dispositions are not inconsequential in many choices and preferences that we make daily, routinely, or occasionally.

If we are going to "do better" in our efforts to assess educational outcomes, there is much to be learned from the experiences of the University System of Georgia (USGA) in establishing systemwide entrance requirements, developmental studies programs, reading and writing tests for sophomores, and varying forms of senior exit examinations. All

facets of these programs are the creatures of statewide policies decided by a single governing board for 34 institutions of public higher education.

Clearly demonstrated by such systemwide requirements is a strong need for the guidance and direction of policy. Also needed is a coherent rationale that will lend credibility and utility to institutions and academic departments that do not view the assessment of educational outcomes from a systemwide perspective. A bit of historical background should help explain.

A historical precedence for systemwide testing has long been established in the University System of Georgia. Early in the University System's history an office of "University Examiner" was set up for the preparation, administration, and scoring of systemwide tests in subject-matter areas. The eventual discontinuance of that office need not refute its importance as a precedence for statewide testing. In 1957 the Board of Regents adopted the SAT as an admission requirement for all units (n=18) of the University System. The SAT continues to be a requirement for entry to units of the University System and is one of the few continuous statewide testing programs in higher education.

In 1974, a program of developmental studies was established whereby each unit (with the exception of the Medical College of Georgia) would offer up to four quarters of remedial education to students not meeting regular admission requirements but showing significant promise for college work. A systemwide test in basic skills was developed for purposes of placement in appropriate coursework. To reduce disruptions in student transfer from institution to institution, a core curriculum had been established earlier and a Regents Testing Program was organized for purposes of convincing senior colleges that graduates of junior colleges could read and write well enough for upper division work. And finally, a mandate was issued whereby each unit of the University System would develop some kind of senior exit test(s) and submit a proposal for such to the Board of Regents for approval.

The experience with systemwide testing in the University System of Georgia suggests that in any concerted efforts to assess educational outcomes, there should be at least three stages of assessment. A curve cannot be plotted with only two points, and simple "before/after differences" in test scores can be misleading. Multi-stage assessment had proven effective in the 1950s at Georgia State University where a comprehensive testing program of academic ability and educational achievement was established at entry, sophomore-junior, and senior exit levels (See McClain & Krueger, 1985 for more recent efforts). Under an injunction in the late 1950s not to discriminate against minority groups in admissions, Georgia State adopted an extensive series of entrance examinations, consisting of the Otis Quick-Scoring Test of Mental Ability (Otis); the Cooperative General Achievement Tests in Social Studies, Natural Sciences, and Mathematics (GAT-Social Studies; GAT-Science; and GAT-Math); the New Cooperative English Test (Coop. English); and the Nelson-Denny Reading Test (N-D Vocabulary; N-D Reading Comprehenssion). The battery of tests were required of all students applying to undergraduate programs. Entering freshmen were required to submit SAT scores in addition to the entrance examinations, and thus, a total of nine standardized test scores were available for each freshman seeking admission.

The Junior Achievement Testing Program consisted of the Cooperative General Culture Tests (in social studies, literature, science, mathematics, and fine arts) and the Cooperative English Expression Test—both tests having been chosen by a faculty committee for their comparability (at an apparent higher level) to tests included in the entrance examinations.

For graduating seniors in the School of Arts and Sciences, the exit exams consisted of the Graduate Record Exam (GRE-Verbal and GRE-Math) and the GRE Advanced Test (GRE-ADV). The value of assessing educational outcomes at three different stages was demonstrated when sufficient data permitted the correlation of entrance, junior achievement, and graduating senior exit tests with each other and with each senior's grade-point-average (GPA). Mean scores showed an expected rise from entry to junior to senior levels and correlation coefficients confirmed

logical expectations about achievement in related areas. The highest correlation (+.46) was found between GRE-ADV and senior GPA; the lowest correlation (-.01) was between mathematical achievement at the freshman level and senior GPA.

The choice of the GRE as a senior exit exam proved to be fortunate. Many arts and sciences seniors were considering graduate work and would have taken the GRE on their own. The tests were administered at institutional expense and thus provided a financial incentive to some students. The requirements of senior exams were published in the college catalog for two years prior to implementing the program, and there was a precedent by the School of Business who had for years required senior exams for business majors. There was also the precedent of the junior achievement tests which were entered as part of the individual's academic record and for which concerted efforts were made to provide knowledge of results. Each student received in writing a fairly detailed explanation of his or her scores and was encouraged to discuss the test results further with a counselor in the Office of Testing and Counseling. Such precautions seem to have prevented what later occurred at the University of Georgia when the GRE was required (without sufficent notice) of graduating seniors. Reading their letters of notification as a requirement to take the test (but not necessarily to pass it), many seniors came to the testing session, signed their names to the test, turned it in, and left. As a result, University of Georgia students may have some of the lowest GRE scores ever recorded.

Student (and institutional) reactions to a senior exit exam mandated by the Board of Regents resulted in the appointment of an ad hoc committee to study ways in which some kind of program assessment could be implemented systemwide. One outcome of the committee's work was the development of a rationale for program assessment (1977). All the difficulties of defining an academic program for purposes of evaluation were acknowledged, and a consistent effort was made to place responsibility at the academic department level. Departments of instruction were free to select, develop, or otherwise devise their own methods of program assessment. Distinctions were made between assessment and evaluation, with the former advocated as the more appropriate term.

Despite the development of a rationale for departmental responsibility in the assessment of academic program outcomes, the study committee could not surpass the impeccable logic of Chancellor George L. Simpson, who expressed the need to assess educational outcomes this way:

> If the faculty members of each academic department would prepare a comprehensive examination of program objectives and expected outcomes, and if they would administer the exam to graduating seniors, grade the exam, and then discuss the exam with each student, faculty members would then know if the program is accomplishing its purpose—and they certainly would know if graduating seniors could read and write sufficiently.

The published rationale did have the value of demonstrating to academic department heads the many sources of assistance available to them in their assessment of program objectives and outcomes. Clearly stated was the possibility of using nationally standardized tests when such tests were appropriate for purposes of program assessment. For example, where a nationally standardized test in a subject matter area could be administered, and where graduating seniors scored above the reported national norm for such a test, it would be difficult to imagine a more direct, more convenient way to satisfy the Board of Regents mandate for program assessment.

Unfortunately for those who would emulate the University System of Georgia's three-tier program of entry, rising junior, and senior exit assessment, there is much to be cautious about! The requirement of such tests is not free of administrative, legal, and ethical entanglements. The developmental studies programs of USGA institutions have been reviewed with no overwhelming evidence that they are successful in teaching basic skills of literacy (i.e. reading, writing, and reasoning as required in traditional college coursework).

The Basic Skills Examination (BSE) has apparently been discontinued in favor of a test yet to be constructed, developed, and validated. And the Regents Testing Program for rising juniors continues to be the object of potential litigation. Students in the public historically black institutions continue to have a higher failure rate than students in predominantly white institutions—and the test (and the manner in which the writing section is graded) is frequently regarded as culturally biased. The use of the test raises many questions of fairness to students with cultural differences in a statewide system of 34 diverse institutions. And there remains an incredible lack of information about the general or specific outcomes the test might assess—and how these are related to educational achievement at junior and senior levels of educational achievement. Periodically the posting of passing rates by institution (usually in the Atlanta newspapers) raises questions about the test's validity and utility.

In closing, it should be emphasized that systematic, objective, valid, reliable, fair, and creditable measures of educational outcomes are rare. Despite shopping lists that have been publicized by advocates of outcomes assessment (Adelman, 1986; Ewell, 1985), there are very few "ready-made" methods of assessment available to the college with little or no measurement, assessment, or evaluation expertise in its administrative staff and faculty. Neither would there seem to be much agreement about the results institutional leaders should expect from assessment. It is distinctly possible to assess student mastery of terms, concepts, and general principles; it is often possible to measure student comprehension, analysis, and interpretation of the structure and substantive issues of their respective major fields; and it is even possible to assess in various ways student competencies in decision making and problem solving—BUT many institutions do not have the resources and capabilities to do so, and their leaders do not have educationally sound expectations as to what they would find if educational outcomes in their institutions were carefully assessed.

The gist of this paper thus might be: while awaiting more adequate concepts and theories of educational achievement, and while anticipating better theories and methods of assessment, it would behoove college faculties and administrative staffs to measure what they can, assess what they must, and evaluate only with the greatest of care!

POSTSCRIPT

The decision to publish this commentary has been spurred by numerous events taking place since the fall of 1987. State Higher Education Executive Officers (SHEEO, 1987) have issued a policy statement in which the assessment of achievement in general education is recommended for all undergraduate-degree-granting institutions. Also recommended are: (a) the assessment of basic skills for all entering freshmen, and (b) the use of licensure and certification exams as an appropriate measure for judging program and institutional quality. The American Council on Education (ACE) and the American Association for Higher Education (AAHE) have issued a joint statement on assessment in which perspectives are provided presidents and chief academic officers (Rossmann and El-Ehawas, 1987).

To document "the assessment movement," ACE has surveyed the nation's institutions and found that three out of four colleges are considering assessment as a means to institutional effectiveness; the majority of academic administrators expect assessment efforts to increase, and they support assessment efforts that are tied to instructional improvement (El-Khawas, 1987). The Education Commission of the States (ECS) has surveyed state boards of higher education and found that two-thirds of the states have initiated efforts involving "assessment" of one kind or another (Boyer, et al., 1987).

National testing agencies are now moving into the assessment "market" and developing instruments that will serve the assessment needs of colleges. The College Board (1987) has developed a CLEP Education Assessment Series to assess the gains in general education during the freshmen and sophomore years. Educational Testing Service (ETS, 1987) is developing a new assessment service in which basic skills are related to areas of general education. And the American College Testing Program (ACT) has announced its Collegiate Assessment of Academic Proficiency program in which reading, writing, math, and critical thinking can be assessed.

In 1988 there is no doubt that assessment is much upon us! The groundswell is a function of public demands for accountability, institutional needs for accreditation, and national expectations concerning the improvement of undergraduate education. It all suggests that once again, we live in "interesting times."

REFERENCES

Academic Degree Program Assessment in the University System of Georgia: A Tentative Statement of Rationale, Purposes, and Procedures. Atlanta: University System of Georgia, 1977.

Academic Preparation for College: What Students Need to Know and be Able to Do. New York: The College Board, 1983.

Adler, Mortimer J. The Paideia Proposal: An Educational Manifesto. New York: Macmillan, 1982.

American College Testing Program. "CAAP: Collegiate Assessment of Academic Proficiency." Iowa City: ACT, 1987.

Bowen, Howard R. Investment in Learning. San Francisco: Jossey-Bass, 1977.

Boyer, Carol, et al. Assessment and Outcomes Measurement: A View From The States. Denver: ECS, 1987.

College Board. Outcomes Assessment in Higher Education: College Administrators Workshop Guide. New York: CEEB, 1987.

Educational Testing Service. "The Academic Profile." Princeton: ETS, 1987.

El-Khawas, Elaine. Campus Trends, 1987. Washington, D.C.: American Council on Education, 1987.

Ewell, Peter (Ed.). Assessing Educational Outcomes. New Directions in Institutional Research No. 47. San Francisco: Jossey-Bass, 1985.

Harris, John. "Assessing Outcomes in Higher Education." In Adelman, C. (Ed.) Assessment In American Higher Education. Washington, D. C.: U.S. Department of Education, 1986.

McClain, Charles J. and Krueger, Darrell W. "Using Outcomes Assessment: A Case Study in Institutional Change." In Ewell, Peter (Ed.) Assessing Educational Outcomes. San Francisco: Jossey-Bass, 1985.

Preparing for College: Essential Courses and Skills. Atlanta: University System of Georgia, 1985.

Rossmann, Jack and El-Khawas, Elaine. Thinking About Assessment: Perspectives for Presidents and Chief Academic Officers. Washington, D.C.: American Council on Education, 1987.

State Higher Education Executive Officers. "A Statement of Policy by the State Higher Education Executive Officers on Program and Institutional Assessment." Denver: SHEEO, 1987.

IHE Newsletter
February 1988
Athens, Georgia 30602

Cost: $310.00/Quantity 1,000
UPD 8897/2-88